CS 5306
INFO 5306:
Crowdsourcing and Human Computation

Lecture 17

10/31/17

Haym Hirsh

TECH

This AI came up with some hilarious Halloween costumes

By Lauren Tousignant

October 30, 2017 | 3:20pm

TRENDING NOW
ON NYPOST.COM

Delta passengers busted having oral sex after meeting on plane

TECH

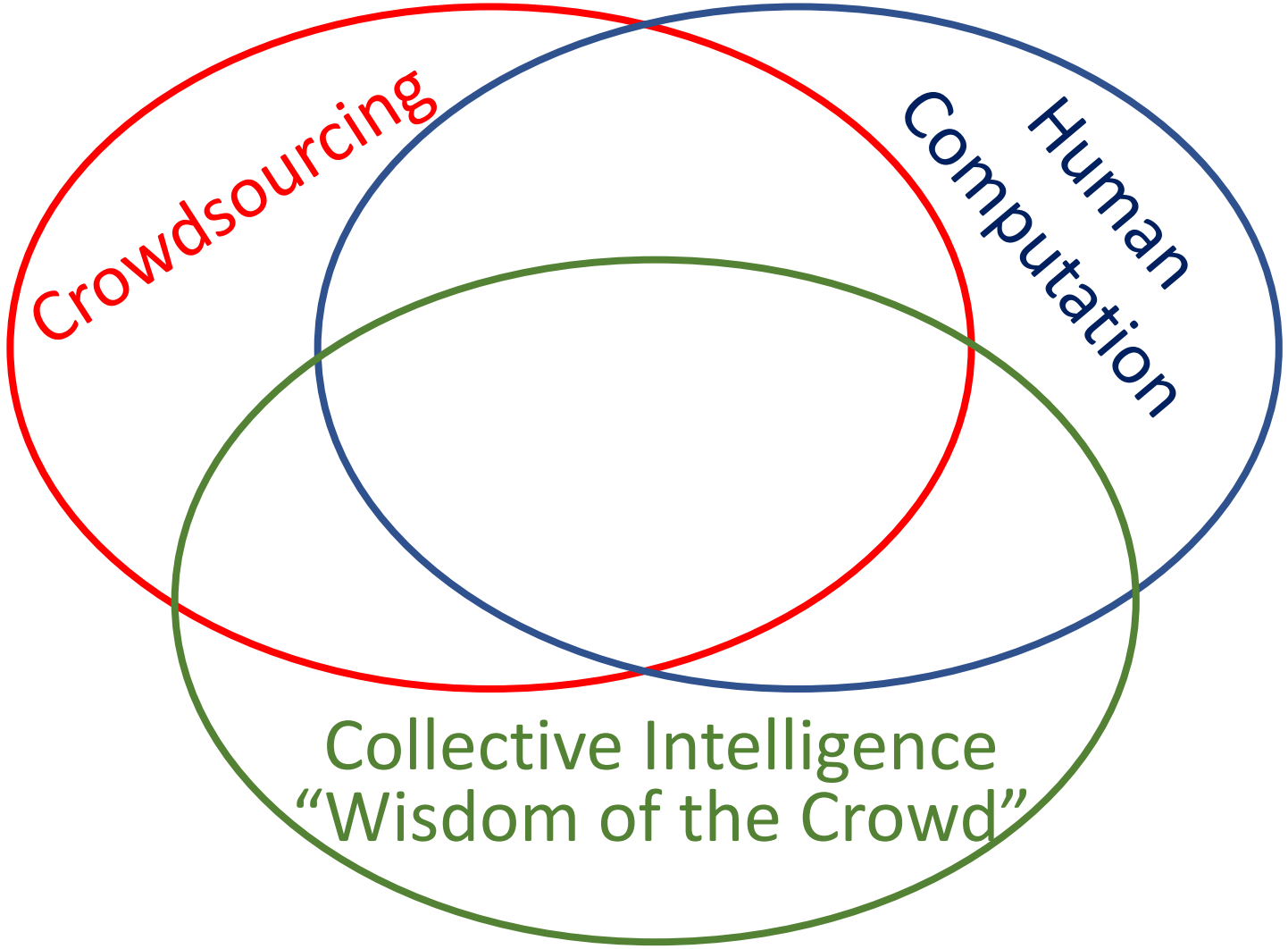
This AI cam hilarious Ha

By Lauren Tousignant

To set up the Halloween generator, Shane entered 4,500 costume names into the network after crowdsourcing from the internet. The AI then took that data and tried to learn patterns and combine words.

TRENDING NOW
ON NYPOST.COM

Delta passengers busted having oral sex after meeting on plane


Crowdsourcing
Human Computation
Collective Intelligence
“Wisdom of the Crowd”

Collective Intelligence

“Collective intelligence is a shared or group intelligence that emerges from the collaboration and competition of many individuals and appears in consensus decision making in bacteria, animals, humans and computer networks” - Wikipedia

Collective Intelligence

- Examples:
  - Animals:
    - Birds
    - Fish
    - Bees
    - Ants
    - Termites
  - Mold
  - Bacteria

Collective Intelligence

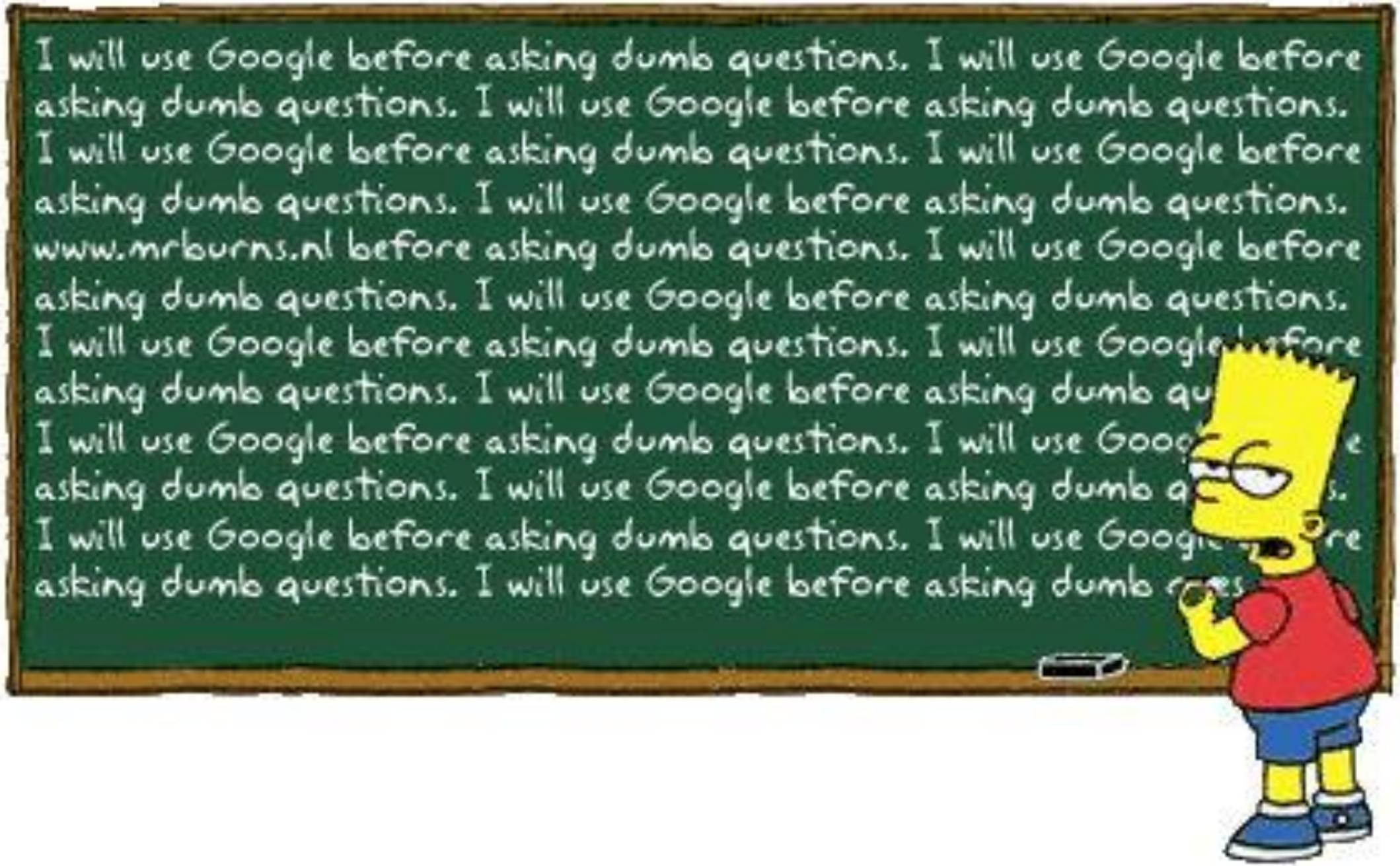
- Examples:
  - Animals:
    - Birds
    - Fish
    - Bees
    - Ants
    - Termites
  - Mold
  - Bacteria

People

Collective Intelligence

- Examples:
  - Animals:
    - Birds
    - Fish
    - Bees
    - Ants
    - Termites
  - Mold
  - Bacteria

People
Online


I will use Google before asking dumb questions. I will use Google before asking dumb questions. I will use Google before asking dumb questions. I will use Google before asking dumb questions. I will use Google before asking dumb questions. I will use Google before asking dumb questions. www.mrburns.nl before asking dumb questions. I will use Google before asking dumb questions. I will use Google before asking dumb questions. I will use Google before asking dumb questions. I will use Google before asking dumb questions. I will use Google before asking dumb questions. I will use Google before asking dumb questions. I will use Google before asking dumb questions. I will use Google before asking dumb questions. I will use Google before asking dumb questions. I will use Google before asking dumb questions. I will use Google before asking dumb questions. I will use Google before asking dumb questions. I will use Google before asking dumb questions.

Google: Pagerank

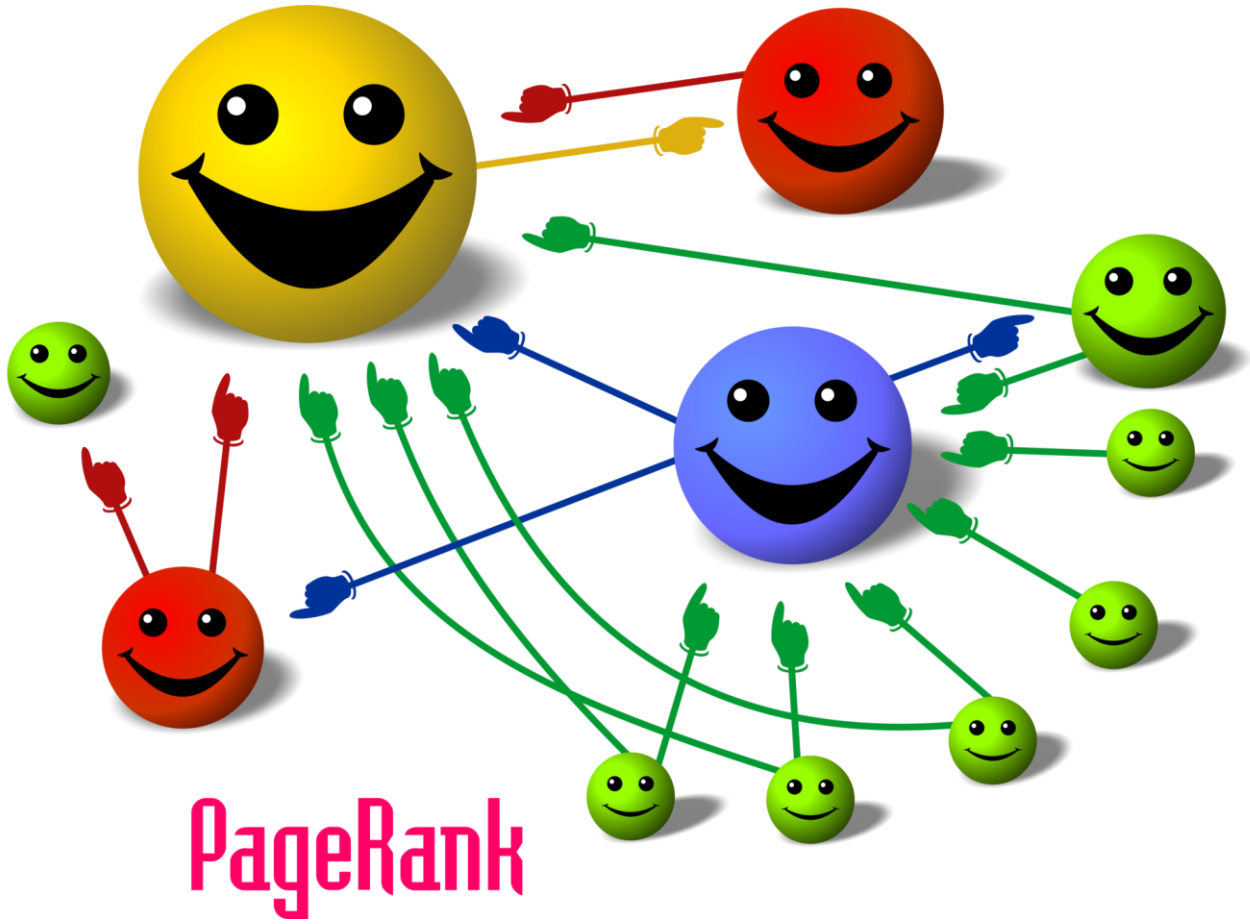
- Pagerank: “The anatomy of a large-scale hypertextual Web search engine”, S. Brin and L. Page. In *Proceedings of the seventh international conference on the World Wide Web*, 1998.

Google: Pagerank

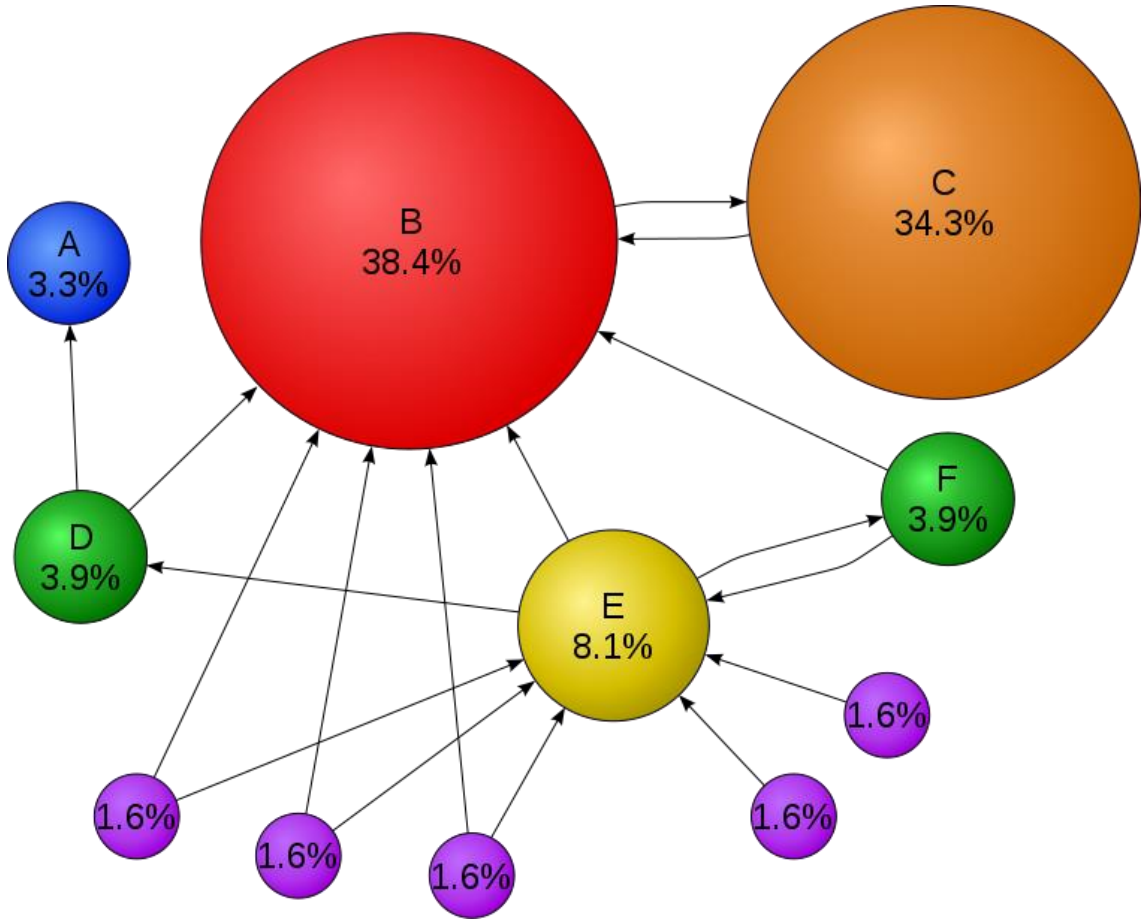
- Pagerank: “The anatomy of a large-scale hypertextual Web search engine”, S. Brin and L. Page. In *Proceedings of the seventh international conference on the World Wide Web*, 1998.

- HITS: “Authoritative sources in a hyperlinked environment”, Kleinberg JM. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms* 1998.

Pagerank



Pagerank



Pagerank

Imagine someone going randomly from page to page according to the popularity of a page ("random surfer model").

For a page V_i:

$$S(V_i) = \sum_{j \in in(V_i)} \frac{S(V_j)}{|out(V_j)|}$$

$S(V_i)$ is proportional to the probability of landing on V_i.

Pagerank

- Problem: No way to leave a “sink”, so you’ll always end up in one
- Solution: $$S(V_i) = (1-d) + d \sum_{j \in In(V_i)} \frac{S(V_j)}{|Out(V_j)|}$$
  - d is the “damping factor” – probability of following a link versus taking a random jump somewhere else

Pagerank

- Solving $$S(V_i) = (1-d) + d \sum_{j \in In(V_i)} \frac{S(V_j)}{|Out(V_j)|}$$

- Do a random walk
  - Start with equal probabilities for all nodes
  - Start at a random node
  - Take a step according to the probabilities of the edges (or make a random leap)
  - Repeat

- This converges to the correct S
- It can be solved using known method in computational linear algebra

Google Search Today

- Pagerank is one of 200 factors
  - Site specific
  - Page specific
  - User specific]